

# **A Project Report**

**On**

**“Neurological Disease Probability Prediction Using Voice Characteristics”**

**Submitted in partial fulfilment for the requirements of**

**B.E. (CSE) VII Semester Project: PART - I**

**in**

**Computer Science and Engineering**

**By**

<b>Roll Number</b>	<b>Name</b>	<b>Project Supervisor</b>	<b>Supervisor Signature</b>
160119733182	Arjun Gandotra	Smt. M. Anila	



**Department of Computer Science and Engineering,  
Chaitanya Bharathi Institute Of Technology (Autonomous),  
(Affiliated to Osmania University, Hyderabad)**

**Hyderabad, Telangana (India) - 500075**

**December, 2022**



CHAITANYA BHARATHI  
INSTITUTE OF TECHNOLOGY (A)

Kokapet ( Village), Gandipet, Hyderabad, Telangana-500075. [www.cbit.ac.in](http://www.cbit.ac.in)



ISO Certified  
9001:2015

COMMITTED TO  
RESEARCH,  
INNOVATION AND  
EDUCATION

43  
years

## CERTIFICATE

This is to certify that the case study titled “**Neurological Disease Probability Prediction Using Voice Characteristics**” is the bonafide work carried out by **Arjun Gandotra (160119733182)** of B.E.(CSE) of Chaitanya Bharathi Institute of Technology, Hyderabad, affiliated to Osmania University, Hyderabad, Telangana(India) during the academic year 2021-2022, submitted in partial fulfilment of the requirements for the award of degree Bachelors of Engineering (Computer Science and Engineering) and that the project has not formed the basis for the award previously of any other degree, diploma, fellowship or any other similar title.

**Supervisor:**

Smt. M. Anila

Assistant Professor, Dept. of CSE,

CBIT, Hyderabad

# DECLARATION

I, hereby, declare that the case study work entitled “**Neurological Disease Probability Prediction Using Voice Characteristics**” submitted to Chaitanya Bharathi Institute of Technology Hyderabad, is a record of an original work done by **Arjun Gandotra (160119733182)**, and the project has not formed the basis for the award previously of any other degree, diploma, fellowship, or any other similar title.

(Arjun Gandotra)

## ACKNOWLEDGEMENT

I pass my sincerest gratitude to the respected, **Smt. M. Anila** for their grand guidance. I, hereby, thank her for providing us with this great opportunity to submit my report on this prestigious topic. I request her blessings for all my future endeavours. I would also like to thank the HOD, **Dr. Y. Rama Devi**, for providing me with the best resources possible for the completion of this report. She has been a great help.

Furthermore, I would like thank Dr. K Sagar, Dr. Rupesh Mishra, and Smt. S. Durga Devi for their extraordinary cooperation and invaluable feedback.

I would like to extend my most cordial gratitude to my parents, without whose support, I wouldn't have been able to achieve the level of intelligence that has, in all earnestness, rendered me fit for this great and prestigious opportunity. I would also like to thank all the personnel related to the research papers for all the work and time they have spent on this.

# INDEX

## LIST OF CONTENTS

S. No.	Contents	Page Number
1	Abstract	6
2	Introduction <b>1.1</b> Problem Definition <b>1.2</b> Scope of the Project	7
3	Literature Survey and Technologies Used	8
4	Existing System	10
5	Design of the Proposed System	11
6	Conclusion	12
7	References	13

## LIST OF ILLUSTRATIONS

Image No.	Description	Page Number
1	Transcript Comparison	8
2	Feature Graphs	9
3	Pipeline Approach	9
4	End-to-End Approach	9
5	Existing System	11
6	Proposed System	12

## **Abstract**

Nearly 1 billion people in the world suffer from some neurological disorder. Majority of people suffer from Alzheimer's Disease or Parkinson's Disease. These neurological diseases are basically an anomaly of the central nervous system. The difficulty in the diagnosis is that every patient has their own set of symptoms. But the most common symptom among them all is variation in their voice. So, data extracted from the voice samples of these patients can be used to predict the probability of them having a neurological disease using a neural network.

This project aims to construct such a model to employ vocal characteristics to classify and predict the probability of a person having a neurological disease.

# **Chapter 1**

## **INTRODUCTION**

### **1.1 PROBLEM DEFINITION**

There are many neurological diseases known, such as Alzheimer's Disease or Parkinson's Disease. Around 1 billion people in the world suffer from a neurological disease. This project will be focusing on Parkinson's Disease. Almost every PD patient has their own set of symptoms which differ greatly from others, but a few basic symptoms are common in mostly all of them, namely resting tremors, stiffness, and slowness of movement. These resting tremors and stiffness contribute to another common symptom that is voice tremors or variation in speech. This variation in speech is the basis for this project. This project aims to construct a model to employ vocal characteristics to classify and predict the probability of a person having a neurological disease like PD. Actually, there is no clinical test present at the moment to properly classify a person as a PD patient. So, this system aims to assist the doctors with diagnosis.

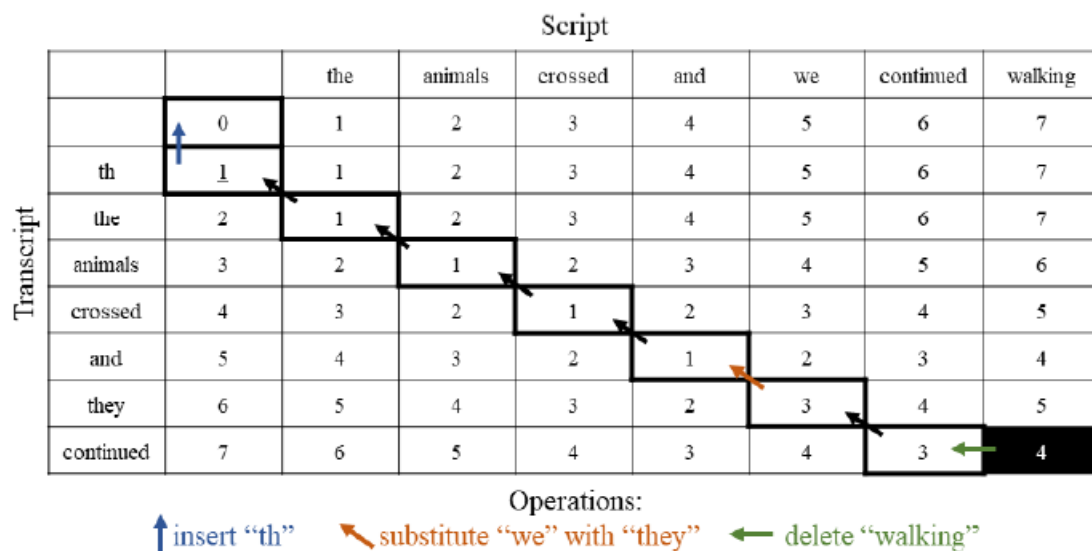
### **1.2 SCOPE OF THE PROJECT**

- Decreasing the noise of the input data.
- Major feature extraction to best optimize the neural network.
- Training a Recurrent Neural Network having at least 3 hidden layers.
- Choosing the activation function for best performance.
- Hyper-parameter tuning using various techniques to best fit the data.

## Chapter 2

### LITERATURE SURVEY

People with neurological diseases like PD also suffer from Cognitive Impairment(CI). The cognitive impairment causes speech disfluencies and errors. The author of this paper [3] analyses the changes in read patterns caused by errors and disfluencies as a function of CI. They start by analysing errors and disfluencies in manual transcripts with annotations specifying where and which errors and disfluencies occur. Then, they detect the same errors and disfluencies from manual transcripts without annotations and couple that detection with Automatic Speech Recognition(ASR) to generate transcripts and detect errors and disfluencies. They found that these features can be used in a regression framework to predict MoCA Scores. Thus, they hypothesize that these features will be sensitive to cognitive decline for people without PD, but the specific error and disfluency patterns may differ.



*Figure 1: Transcript Comparison*

The articulatory and phonatory aspects of speech can be used to assess neurological diseases better [1]. They conclude that although many studies use the acoustic features such as jitter and shimmer, the usefulness of phonatory approaches in the diagnosis of neurological disorders is not apparent yet. So, these features cannot help in distinguishing between patients with neurodegenerative diseases and controls. Some new studies indicate that the phonatory features like noise can be used to differentiate between the two but are more suitable for automatic detection or assessment approaches rather than being employed as single biomarkers. Regarding the articulatory approaches, VSA features family has also shown useful differentiation between patients and controls as it basically characterizes a reduction of the articulatory movements' extension while producing vowels in



connected speech. But we have to keep in mind that all these patients are trying to speak in English.

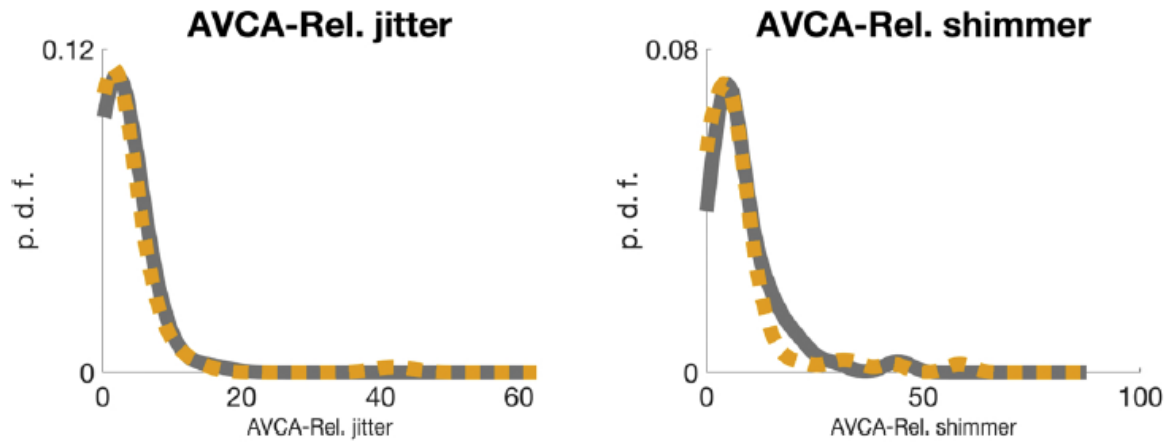


Figure 3: Feature Graphs

This study [5] uses two classifier architectures: traditional pipeline approach and end-to-end approach. In the pipeline approach, they use glottal features to extract the articulatory and phonatory features. A similar approach was used in the paper [1]. Here [5], they use the iterative adaptive inverse filtering (IAIF) and quasi-closed phase (QCP) glottal inverse filtering methods. The end-to-end approach uses deep learning models which were trained using both raw speech waveforms and raw voice source waveforms. They used the PC-GITA speech database. Apart from the speech waveform, the usage of the glottal source signal as a raw waveform in end-to-end systems is justified because the voice pathologies affect the vocal folds. The accuracy of the SVM-based pipeline method was improved from 65% to 67% when the baseline features were merged. They extracted the baseline features using the NeuroSpeech toolkit. They also concluded that their model could be extended to predict the neurological state of PD patients as well.



Figure 2: Pipeline Approach

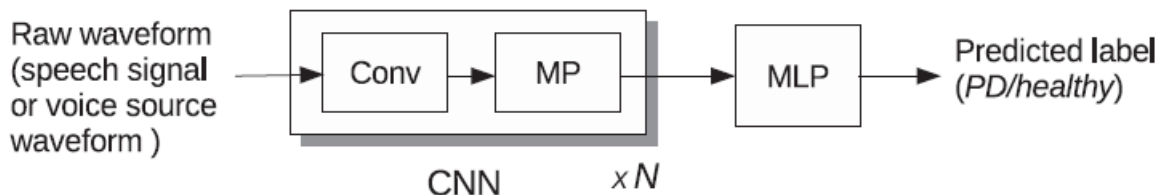


Figure 4: End-to-End Approach

The author of this paper [2] studied the performance of neurological disease detection systems. The performance of such a system degrades when there is an acoustic mismatch between training and operating conditions caused

mainly by degradation in test signals. They try to regulate the quality of the voice signals(recordings) using two different approaches by identifying the presence and type of short-term and long-term degradations and protocol violations. They use the mPower Mobile PD data set under different degradation conditions to test their approaches. Basically, they want to increase/control the quality of the voice signals in order to optimize the vocal characteristics and increase the performance of these detection systems.

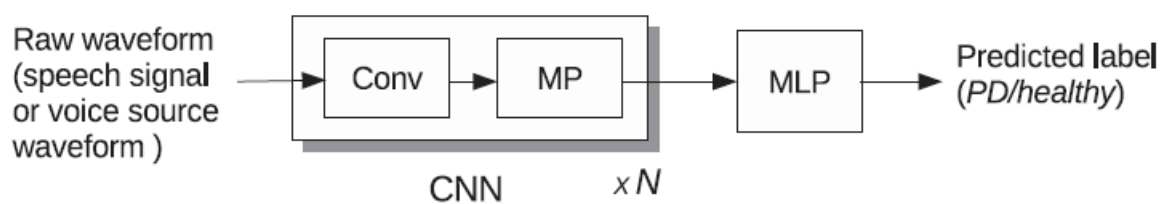
**TECHNOLOGIES USED:**

- Python Programming.
- TensorFlow and Keras Libraries.
- t-SNE.
- Recurrent Neural Network.
- Comparative Analysis.
- Transfer Learning.

## Chapter 3

### EXISTING SYSTEM

A CNN trained with raw waveform image processing to label a patient as having Parkinson's Disease or a healthy control. It uses the waveforms of the vocal features like frequency as input to a CNN model. In this, Max Pooling is done and the output of the Max Pool is given as input to a multi-layer perceptron which is nothing but an ANN with multiple fully connected layers. The output of this model is a label predicting whether a person has PD or a healthy control.



*Figure 5: Existing Methodology*

## Chapter 4

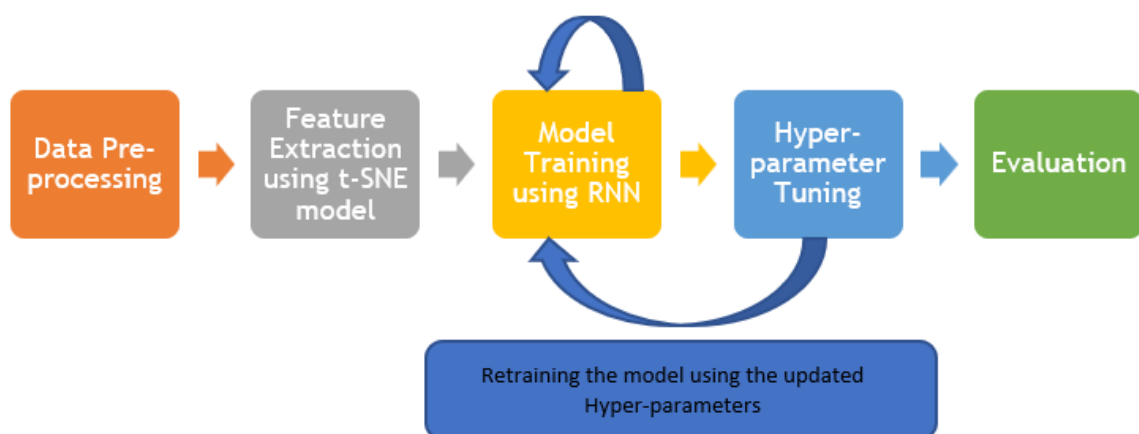
### Design of the Proposed Model

Considering all the base papers, it can be seen that none of the papers use RNN as the main algorithm, so this project will focus on that to get a new system:

Steps involved are:

1. The dataset contains vocal characteristics which being glottal features will have a lot of noise, so some pre-processing will be done to decrease that noise.
2. Feature step using the t-SNE model will be done as it has proven to be effective in the high dimensionality datasets and provides a high accuracy for classification. The extracted features will be split into train and test sets.
3. Training a Recurrent Neural Network, using the keras library for the hidden layers.
4. At each time step, we use the output of the last step as the input to the next, applying an activation function like SoftMax or ReLU or tanh. At the hidden layer, the ReLU function will be used as it provides the same benefits as Sigmoid function but with better performance. At the output layer, the SoftMax function would be preferred as its probabilities are interrelated. They sum up to one. So, if the likelihood of one class increases, the other has to decrease by an equal amount. Thus producing better probability prediction.
5. Afterwards hyper-parameter tuning will be done to enhance the results.
6. To avoid overfitting, the technique of dropout can be utilised.
7. After all this, the evaluation metrics like accuracy, precision and recall will be used to evaluate the model.

### Proposed Algorithm



## **Chapter 5**

### **CONCLUSION**

This work proposes a recurrent neural network for vocal characteristic analysis using various optimization models at different steps of the work flow. Using t-SNE model for feature extraction, variable number of hidden layers starting from 3 layers in the neural network, hyper-parameter tuning using comparative analysis with models like Halving Grid Search, and using the various evaluation metrics to better the model.

This model will predict the probability of a patient having PD, and based on that probability, the patient will be classified as having PD or a healthy control. In future, this project can be used a base for other neurological diseases.

## REFERENCES

- [1] Laureano Moro-Velazquez, Jorge A. Gomez-Garcia, Julian D. Arias-Londoño, Najim Dehak, Juan I. Godino-Llorente, “Advances in Parkinson's Disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects”, 2021, ISSN 1746-8094, <https://doi.org/10.1016/j.bspc.2021.102418>.
- [2] Amir Hossein Poorjam, Mathew Shaji Kavalekalam, Liming Shi, Yordan P. Raykov, Jesper Rindom Jensen, Max A. Little, Mads Græsbøll Christensen, “Automatic quality control and enhancement for voice-based remote Parkinson's disease detection”, 2019, arxiv, <https://doi.org/10.48550/arXiv.1905.11785>
- [3] Amrit Romana, John Bandon, Matthew Perez, Stephanie Gutierrez, Richard Richter, Angela Roberts, Emily Mower Provost, “Automatically Detecting Errors and Disfluencies in Read Speech to Predict Cognitive Impairment in People with Parkinson's Disease”, 2021, ISCA, <https://doi.org/10.21437/Interspeech.2021-1694>
- [4] O. Karaman, H. Çakın, A. Alhudhaif, K. Polat, “Robust automated Parkinson disease detection based on voice signals with transfer learning”, *Expert Syst. Appl.* 178 (2021), 115013, <https://doi.org/10.1016/j.eswa.2021.115013>
- [5] N. P. Narendra, B. Schuller and P. Alku, “The Detection of Parkinson's Disease From Speech Using Voice Source Information”, in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1925-1936, 2021, <https://doi.org/10.1109/TASLP.2021.3078364>